When does Information become Knowledge?

H. V. WYATT*

Biology Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20014

By considering a specific example from molecular biology, Dr Wyatt examines the factors that determine the transformation of scientific information into scientific knowledge.

It is generally accepted that molecular biology began with the ~~paper by~~ Avery, MacLeod and McCarty[1] in 1944. The textbooks often link that paper with the equally well known although perhaps seldom read paper published in 1952 by Hershey and Chase[2]. The next year the double helix structure for DNA was proposed by Watson and Crick[3] and molecular biology became established. The history of these first ten years has been viewed through eyes perhaps mellowed by hindsight and wishful thinking but has not been examined in the light of the published work (as opposed to the memoirs) of the time. The fact that Avery was not awarded a Nobel Prize seems strange unless his work was less well known or was less recognized than we might now suppose. In this article I shall examine the information content of Avery's paper, the citations to it and the recognition it received in reviews and books.

Information Sources

Nowadays we may become acquainted with a paper directly or through a number of secondary sources such as *Science Citation Index*, *Current Contents*, keywords in *Biological Abstracts*, through subject groupings in *Biological Abstracts*, *Index Medicus* and more specialized sources like *Genetics Abstracts* (see ref. 4). Although there were fewer aids in 1944, the keys were similar and there were fewer papers to find.

The paper by Avery *et al.* was published in *Journal of Experimental Medicine* from the Rockefeller Institute where Avery worked. It was a famous and respected journal bought usually by medical rather than science libraries, but at that time most geneticists worked in science departments. In 1944, there were thirty-six copies of the journal in libraries in Britain and of these at least twenty and probably more than twenty-six were in medical or veterinary libraries[5]. In universities like Leeds the medical school and its library are separated from the main campus. In London the only copies apart from those in medical libraries were in the libraries of the Chemical Society and the University of London, which was a closed access library. The Second World War would also have exaggerated the usual delays in delivery of the journal outside the United States.

In libraries in the United States itself there were about 125 copies[6] of the *Journal of Experimental Medicine*. Of these at least forty-five were in medical libraries, probably seven in the libraries of pharmaceutical firms and seventy in university libraries. Some of those in university libraries were probably housed in separate medical buildings; there were at least two in agricultural libraries. The records of the journal show that there were about 600 paid and 100 complimentary copies in the United States. Thus in the United States, this journal may have been more widely accessible to geneticists than in other countries.

Title and Conclusion

The title (Table 1) contained few if any key words that would have led to recognition at the time. Transformation was the term used in the various papers on the subject, all published in medical science journals since Griffith's paper[7] in 1928. "Transformation" also described the work of Berry and Dedrick[8] with Shope's virus. These experiments were known to geneticists but were regarded more as curiosities than part of the mainstream of genetics. Similarly, geneticists did not believe that desoxyribonucleic acid was of great importance in heredity. A reader, therefore, who had glanced at the title might well have missed the significance of the paper, and the conclusion (Table 1) would not have enlightened him.

Table 1 The Title, Summary and Conclusion of Avery *et al.*[1]

Avery, O. T., MacLeod, C. M., and McCarty, M. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types. Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from *Pneumococcus* Type III.

SUMMARY

1. From Type III pneumococci a biologically active fraction has been isolated in highly purified form which in exceedingly minute amounts is capable under appropriate cultural conditions of inducing the transformation of unencapsulated R variants of Pneumococcus Type II into fully encapsulated cells of the same specific type as that of the heat-killed microorganisms from which the inducing material was recovered.
2. Methods for the isolation and purification of the active transforming material are described.
3. The data obtained by chemical, enzymatic, and serological analyses together with the results of preliminary studies by electrophoresis, ultracentrifugation, and ultraviolet spectroscopy indicate that, within the limits of the methods, the active fraction contains no demonstrable protein, unbound lipid, or serologically reactive polysaccharide and consists principally, if not solely, of a highly polymerized, viscous form of desoxyribonucleic acid.
4. Evidence is presented that the chemically induced alterations in cellular structure and function are predictable, type-specific, and transmissible in series. The various hypotheses that have been advanced concerning the nature of these changes are reviewed.

CONCLUSION

The evidence presented supports the belief that a nucleic acid of the desoxyribose type is the fundamental unit of the transforming principle of Pneumococcus Type III.

Summary and Discussion

The summary shown in Table 1 was, again, not very provocative—there was no mention of gene, mutation or any terms to link the findings to general genetical ideas. The discussion was more than three pages long and the last half page dealt with the genetic interpretation. Three current views of the nature of the transforming substance were put forward; namely, that it had been likened to a gene, that it was like a virus and that it was a "transmissible mutagen".

Abstracting Journals

Although the significance of the work was not immediately evident from the title and from the paper itself, it could have been made clearer in the abstracts published in *Biological Abstracts* and *Chemical Abstracts* in 1944. In the first the abstract prepared by McCarty was in the "Immunology, General and Bacterial" section. It did not make clear any genetic importance, nor did the abstract in the "Microbiology" section of *Chemical Abstracts*. In the subject index it could

* On leave from the University of Bradford, England.

have been found through "desoxynucleic acid" and "Pneumococcus" in *Chemical Abstracts* and under "nucleic acid" and "Diplococcus" in *Biological Abstracts*. Anyone searching under genetical terms such as acquired characters, crossing over, dissociation, genes, genetics, genotype, heredity, hybrid, mutation, phenotype and variation, would not have tracked down the paper through either abstracting journal. For retrieval purposes, the paper was evidently considered only from the narrower microbiological point of view. In fairness to the abstracting journals, however, the authors did not provide the leads in their paper and McCarty wrote an abstract which ignored the genetic aspects. It would have been a bold subject editor who would have filed it under "mutation" or "hybrid". It is in such a case that the weakness of the abstracting and indexing journals is most clearly seen. Thus an eager searcher in the secondary journals would have had difficulty in finding the paper and the abstract itself would not have been very revealing.

Professor Coburn has written recently[9] of his friendship with Avery and states that Avery was well aware of the implications of his discovery. This is not so evident from Coburn's narrative. It was eight weeks after his conversation with Avery that Coburn himself realized the implications, which he did not enumerate in his letter to Avery. In a long letter written on May 17, 1943, to his brother, Avery clearly says that his transforming principle (DNA) "sounds like a virus but may be a gene". This letter is reproduced in full by Dunn[10] and by Hotchkiss[11]. Sir MacFarlane Burnet, on a visit to the United States from Australia, spoke with Avery at the Rockefeller Institute a month after the paper was submitted for publication. Burnet wrote to his wife in December 1943, "Avery . . . has just made an extremely exciting discovery, which, put rather crudely, is nothing less than the isolation of a pure gene in the form of desoxyribonucleic acid"[12].

We may assume then that the lack of a clear statement of the importance of his discovery in his paper was due to Avery's "constant modesty and deep humility . . . (whose) high regard for the printed word deterred him from theorizing in print"[9]. His modesty also ensured that the citations contained none which would have firmly linked it to genetics, a reference, for example, to Beadle and Tatum's paper[13] of 1941. Whatever Avery thought of his work, he intended it to be found, seen and read by those interested in pneumococci, not genetics.

Nevertheless the news spread fast. Burnet records that he next visited Cold Spring Harbor and then Beadle and Tatum at CalTech. Demerec, the director of the Cold Spring Harbor Laboratory, was a member of the phage group and another member, Luria, has said that he was well aware of Avery's work before publication[14]. Neither the geneticists nor the phage group, however, appeared to be greatly influenced by the news. But Erwin Chargaff of nearby Columbia University in New York was moved to take up the study of nucleic acids[15,16] and his later work provided an important clue for Watson and Crick.

In an address to the Royal Society in November 1945 on "The Gene", Muller[17] devoted about 3 per cent of his time to "possible roles of the nucleic acid". This brief review, summarizing Avery's work, listed a number of interpretations but left open the question of whether the specificity resided in the nucleic acid polymer or in an associated protein.

Cold Spring Harbor

The 1946 symposium, the first since 1942, was on heredity and variation in microorganisms. There were twenty-seven papers including one by McCarty, Taylor and Avery[18]. Of the 136 participants from six countries, eleven came from the Rockefeller Institute. Six of those present later became Nobel Laureates. One of the four attending from Britain was Pirie from Rothamsted Agricultural Research Station which had the *Journal of Experimental Medicine* in its library. Four of the speakers, Dienes, Hershey, Kidd and Luria, made six references to the three papers by Avery and his co-workers in the *Journal of Experimental Medicine*[1,19,20] and Anderson referred to the presentation at the symposium. Spiegelman made three general remarks about Avery's work without giving citations.

It is therefore surprising that the next symposium, in 1947, on nucleic acids and nucleoproteins was not attended by Avery or MacLeod or McCarty. Chargaff mentioned that "the epochal experiments by Avery and his associates have emphasized the very important role by some bacterial nucleic acids in the determination of inherited synthesizing ability"[15]. The next relevant symposium was in 1951, on genes and mutations. Nucleic acids and DNA were not mentioned except in discussions, until pages 445 to 460 when there were two papers on transformation. Here it was clear that the authors were thinking of the DNA as synonymous with the gene.

Even at the 1953 symposium when Watson and Crick presented a paper on DNA there was little mention of Avery and his associates. There were four groups of biologists who were or should have been interested in Avery's findings—the biochemists, the geneticists, the microbiologists and the phage group. Their various paths came together with the Watson–Crick structure for DNA, as told by Asimov[21].

Biochemists

Chargaff, inspired by Avery's work, made many significant researches into the structures of nucleic acids. The most important was the overthrow of the tetranucleotide theory of Levene. Based on the equimolecular proportions of the four bases, this model proposed that the DNA structure was a sugar-phosphate backbone with alternating purines and pyrimidines in groups of four. Such a structure could have little specificity, especially compared with the proteins known to be associated with chromatin. Chargaff and his associates showed that the DNA is characteristic for each species and that the ratios of A : T and C : G were nearly equal but that the ratio of A : T–C : G varied. This made DNA a much more interesting molecule[15] and provided a vital clue for Watson[22].

Geneticists

It is not surprising that the geneticists were little interested in DNA until Chargaff had shown the tetranucleotide theory to be misleading. The evidence came from one hereditable change in one organism and that organism was a bacterium, a class widely believed to have very different genetic properties from other classes of organisms. It was several years before other examples of transformation were shown in *Pneumococcus*, *E. coli* and *Haemophilus* (see ref. 11). For the classical geneticists, transformation was not a system which could be used in their experiments: it was therefore difficult for them to integrate transformation into their ideas. Beadle records[23] that the one-gene–one-enzyme hypothesis was still largely unaccepted in 1951. Thus a key experiment was described as a curiosity in the textbooks (see below) and ignored by the scientists in general but seems to have been kept to the fore largely in semi-popular lectures.

In a mammoth ninety page summary of the symposium on phosphorus metabolism held in 1951, however, Bentley Glass devoted three pages to DNA in bacterial transformation[24]. He wrote: "Transforming agent, in all likelihood DNA, is akin to if not identical with the genetic units. This is the strongest support for the view advanced by Zamenhof and discussed earlier, the view that the genes of organisms are to be found not in the proteins but in the myriad forms of the DNA itself."

About 150 people attended the meetings: Glass, a geneticist, only Herriot from the phage group, Hotchkiss, a microbiologist (who gave a very good account of transformation), Chargaff and Zamenhof who spoke on DNA; the rest were chiefly biochemists.

Phage Group

The phage group, as Luria records, knew of the work before publication but apparently did not act upon it even though one explanation was that transformation was accomplished by a phage, not DNA. One reason was that Luria had proposed a theory which demanded a phage genetic material containing little DNA[11]. In 1951, however, both Herriott (who had attended the meeting on phosphorus metabolism) and Northrop were suggesting that the DNA of phage might be like the transforming principle. Watson, who had been Luria's student, and Maaløe[25] had noticed that much of the labelled parental phage protein remained attached to cellular debris. The next year Hershey and Chase[2], as a result of their famous experiment, plainly suggested that protein had no function in phage multiplication and that DNA had. Their experiments showed that about 20% of the protein and about 75% of the DNA could not be removed from the infected bacteria, 80% of which were viable. Other experiments showed that less than 1% of the protein from parental phage was transmitted to progeny phage. Detection of proteins was limited to those containing sulphur (labelled with ^{35}S). Thus, as the authors pointed out, "whether sulphur-free material other than DNA enters the cell has not been determined". The authors' interpretation of the experiments, as given in detail in the results, discussion and summary, have mostly been vindicated. We now know, however, that essential proteins are injected with the DNA. Moreover a non-conservative replicating mechanism would mean that there would be no transmission of the radioisotope from the parent to the progeny. What is surprising is that the problem of contamination, which had been cited so often to criticize Avery's experiments, was not raised in the context of these phage experiments.

Summaries Compared

It is particularly interesting to compare the summaries of these two papers: Avery's has its non-committal résumé of the facts whereas Hershey and Chase's has its mixture of fact and interpretation—"8. *The sulphur-containing protein of resting* phage particles is confined to a protective coat that is responsible for the adsorption to bacteria, and functions as an instrument for the injection of the phage DNA into the cell. This protein has no function in the growth of intracellular phage. The DNA has some function." At the 1953 Cold Spring Harbor Symposium on viruses, Hershey discussed the limitations of the experiment.

Luria, who was to have been a speaker at a meeting of the Society for General Microbiology in the spring of 1952 at Oxford, England, was denied a passport by the State Department and Watson gave a talk instead. "Almost no one in the audience of over four hundred microbiologists seemed interested as I read long sections of Hershey's letter"[22] describing the famous experiment. Unfortunately we do not know their interest in the work of Avery or Beadle and Tatum.

The Microbiologists

The recipients of the Eli Lilly award in bacteriology and immunology have included M. McCarty (chemical nature and biological specificity of the substance inducing transformation of pneumococcal types, 1946, ref. 26); S. S. Cohen (synthesis of nucleic acid by virus-infected bacteria, 1951, ref. 27) and J. O. Lampen (metabolism of nucleic acid components in bacteria, 1952, ref. 28). In spite of this interest in nucleic acids, other reviewers in *Bacteriological Reviews* were less than convinced until 1952 of the role of DNA in heredity. In a forty page review of recent advances in bacterial genetics[29], Luria devoted a page and a half to Avery and to later work on *E. coli* in a section on specific induction of mutations. He wrote "... prove biological specificity of nucleoproteins can be carried out not only in the protein but also in the nucleic acid moiety".

In a review of bacterial dissociation, W. Braun[30] devoted two lines to Avery and his co-workers. Cohen said that "evidence that DNA plays a genetic role ... while considerable and suggestive is at least indirect"[27]; he did not cite Avery. A review of bacterial transformations by Austrian[31] was, however, more positive and he mentioned that protein contamination of the transforming DNA seemed to be ruled out by the experiments of Hotchkiss[11]—an original objection voiced by Mirsky at the 1947 Cold Spring Harbor symposium. He continued by saying that "a growing body of increasingly precise evidence gives continuing support to the view that DNA is indeed the biochemical determinant of inheritable characters. Much indirect evidence supports the view that DNA is concerned with genetic mechanisms in all living forms".

Mirsky devoted a page to Avery's work in a report of a meeting held in 1950 (ref. 32), but nearly all of this was devoted to arguing that transformation was due to a protein impurity or a phage particle. He seems to have been unaware of Hotchkiss's work in the same institute.

Thus, by 1952, there were small numbers of influential people who had accepted the idea that DNA might carry genetic information. The 1951 Cold Spring Harbor symposium on genes and mutations contained only two papers, by Ephrussi-Taylor and Hotchkiss, specifically on transformation (DNA). But the discussions were more revealing. Lederberg asked "if it were possible that Hotchkiss was dealing with a suspension of chromosomes or fragments consisting primarily of DNA—this would permit crossing-over of the gene". Szybalski mentioned "Muller's suggestion that whole 'free chromosomes' or parts thereof are involved in transformation". Altenburg, however, asked if "it could be a contaminating virus with the DNA".

Textbooks of the period were less satisfactory. For example, Srb and Owen[33] wrote a good textbook in 1952. The index contained the terms *Pneumococcus* and transforming principle, but no mention was made of Avery. Less than one page was devoted to the pneumococcal experiments and transforming principle which "appears to be a nucleic acid" (page 248). But in the index, none of the words desoxyribose, DNA, nucleic acids, nucleoproteins (or RNA) contains an entry for that page. In the two pages on nucleoproteins it is mentioned that the nucleic acid of the "transforming principle" of pneumococcus is of the DNA class. The section on DNA mentions only nucleotides without detailing A, T, C and G.

Advances in Genetics, an annual which began in 1947, contains no reference to Avery's work until 1955 when it was mentioned twice in a review of microbial genetics by Bryson and Szybalski. Hershey wrote a review of inheritance in bacteriophage in 1953, mentioning his own experiments of 1952. Despite the use of the words desoxypentose nucleic acid and nucleic acid in the latter review and the Watson and Crick paper in 1953, the words DNA, desoxyribonucleic acid, RNA or desoxypentose acid do not appear in the indexes to the first eight volumes (including 1956) and the only entry under nucleic acids alludes to the effect of radiation on them. Evidently the indexers did not consider that these compounds were worth indexing even when their contributors used them. Avery's work was, however, quoted outside these areas. In the 1945 *Annual Review of Physiology*, Wright[34] devoted half a page of a 30 page review of the physiological aspects of genetics to Avery, with the rather ambiguous comment that "the great possible significance of this observation in the interpretation of the role of the nucleic acids of chromosomes and of other self duplicating entities is obvious".

Semi-popular Literature

In addition to the recognized scientific journals, there is a large amount of semi-popular literature which is peripheral to the mainstream. It is, however, difficult to judge its effect either in the long or the short term. *American Scientist* carried at least two articles which emphasized the importance of Avery's work. In 1945 Hutchinson[35] devoted a page to a discussion of Avery's recent paper: "It has been likened to gene or virus. An extremely fundamental contribution to all biological sciences." In 1948 Beadle cited Avery in a six page lecture on "genes and biological enigmas": "Pneumococcus

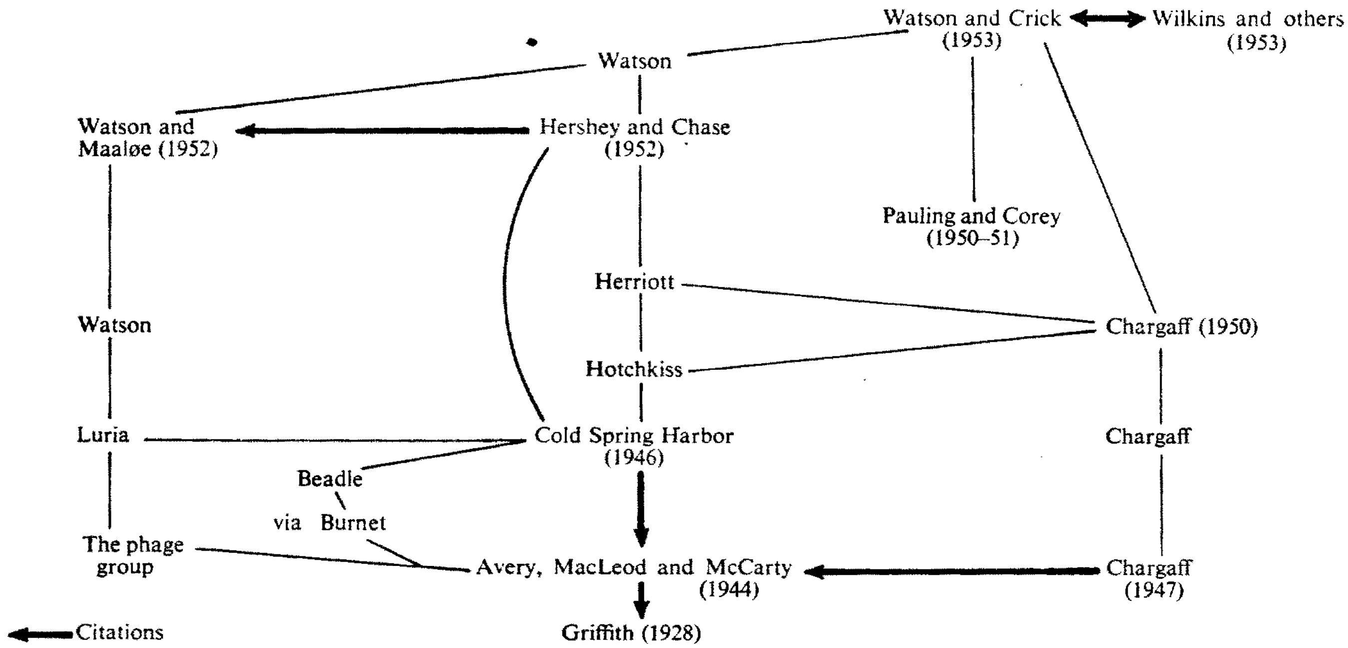

Fig. 1 Simplified picture of the influence, information and citation networks from Avery to Watson and Crick.

type transformations which appear to be guided in specific ways by highly polymerized nucleic acids may well represent the first success in transmitting genes in predetermined ways". Journals like this may influence more students and younger workers than do the primary journals or review series. Visiting lecturers to the campuses, like Beadle, may be more influential for change than the textbooks. The habitual constraints and form of the usual scientific paper become reflected in the textbooks and reviews. It seems that it is only outside the "normal" scientific serials and in discussions, whether public or private, that a freer and bolder view can be offered for debate.

Citation Networks

The invisible college of commuting scientists who visit and correspond with each other is well illustrated in this article and in Watson's book. The various influences are shown in Fig. 1. What is perhaps surprising is the lack of direct citation to the key papers. Thus, Watson and Crick[3] did not cite Hershey and Chase[2] or Chargaff[15] or Avery, MacLeod and McCarty[1]. Similarly, Hershey and Chase[2] did not cite Avery's work[1]. A little later Gierer and Schramm[37], who showed that the RNA is the infective (genetic) portion of tobacco mosaic virus (TMV), cited Hershey and Chase but not Avery and his associates. Garfield, Sher and Torpie[38] made a study of the citation data of the discovery of the genetic code using Asimov's book[21] as the source. Although some of the citation links were good (for example, citations by Avery and his colleagues of the work of Griffith and other workers on transformation) the citation links elsewhere were few. Nor do recent scientists cite Avery: from 1966 to 1969 there was an annual average of seventeen citations to the 1944 paper and an average of four to the other three papers by Avery and his co-workers. Several such citations will have been in articles such as this one.

Information and Knowledge

There is no reason to suppose that with all our information aids today, a similar paper would be any better recognized than Avery's—indeed among so many more papers it might fare worse. Both Beadle[23] and Stent[14] have suggested that new information can be assimilated only when it can be fitted without too much difficulty into accepted ideas. As information it is unrecognized until it is transformed into knowledge.

1 Avery, O. T., MacLeod, C. M., and McCarty, M., *J. Exp. Med.*, **79**, 137 (1944).
2 Hershey, A. D., and Chase, M., *J. Gen. Phys.*, **36**, 39 (1952).
3 Watson, J. D., and Crick, F. H. C., *Nature*, **171**, 737 (1953).
4 Wyatt, H. V., in *Use of Biological Literature*, second ed. (edit. by Bottle, R. T., and Wyatt, H. V.) (Butterworth, 1971).
5 *World List of Scientific Periodicals*, third ed. (Butterworth, 1952).
6 *Chemical Abstracts. List of Periodicals* (American Chemical Society, 1951, 1970).
7 Griffith, F., *J. Hyg.*, **27**, 113 (1928).
8 Berry, G. P., and Dedrick, H. M., *J. Bact.*, **31**, 50 (1936).
9 Coburn, A. F., *Perspect. Biol. Med.*, **12**, 623 (1969).
10 Dunn, L. C., in *Genetic Organisation* (edit. by Caspari, E. W., and Ravin, A. W.), **1**, 61 (Academic Press, New York, 1969).
11 Hotchkiss, R. D., in *Phage and the Origins of Molecular Biology* (edit. by Cairns, J., Stent, G. S., and Watson, J. D.) (Cold Spring Harbor Laboratory, 1966).
12 Burnet, F. M., *Changing Patterns* (Heinemann, London, 1968).
13 Beadle, G. M., and Tatum, E. L., *Proc. Nat. Acad. Sci.*, **27**, 499 (1941).
14 Edsall, J. T., *Science*, **170**, 349 (1970).
15 Chargaff, E., *Essays on Nucleic Acids* (Elsevier, Amsterdam, 1963).
16 Chargaff, E., *Science*, **172**, 637 (1971).
17 Muller, H. J., *Proc. Roy. Soc.*, B, **134**, 1 (1947).
18 McCarty, M., Taylor, H. E., and Avery, O. T., *Cold Spring Harbor Symp. Quant. Biol.*, **11**, 177 (1946).
19 McCarty, M., and Avery, O. T., *J. Exp. Med.*, **83**, 89 (1946).
20 McCarty, M., and Avery, O. T., *J. Exp. Med.*, **83**, 97 (1946).
21 Asimov, I., *The Genetic Code* (New American Library, 1963).
22 Watson, J. D., *The Double Helix* (Athaneum, New York, 1968).
23 Beadle, G. W., in *Phage and the Origins of Molecular Biology* (edit. by Cairns, J., Stent, G. S., and Watson, J. D.) (Cold Spring Harbor Laboratory, 1966).
24 McElroy, W. D., and Glass, B., *Phosphorus Metabolism* (Johns Hopkins Press, 1952).
25 Watson, J. D., and Maaløe, O., cited in ref. 2.
26 McCarty, M., *Bact. Rev.*, **10**, 63 (1946).
27 Cohen, S. S., *Bact. Rev.*, **15**, 131 (1951).
28 Lampen, J. O., *Bact. Rev.*, **16**, 211 (1952).
29 Luria, S. E., *Bact. Rev.*, **11**, 1 (1947).
30 Braun, W., *Bact. Rev.*, **11**, 75 (1947).
31 Austrian, R., *Bact. Rev.*, **16**, 31 (1952).
32 *Genetics in the Twentieth Century* (edit. by Dunn, L. C.) (Macmillan, 1951).
33 Srb, A. M., and Owen, R. D., *General Genetics*, first ed. (Freeman, 1952).
34 Wright, S., *Ann. Rev. Phys.*, **7**, 75 (1945).
35 Hutchinson, G. E., *Amer. Sci.*, **33**, 55 (1945).
36 Beadle, G. W., *Amer. Sci.*, **36**, 69 (1948).
37 Gierer, A., and Schramm, G., *Nature*, **177**, 702 (1956).
38 Garfield, E., Sher, I. H., and Torpie, R. J., *The Use of Citation Data in Writing the History of Science* (Institute for Scientific Information Inc., 1964).